УДК 004.89:004.48

Д.В. Олейник, В.И. Шинкаренко
Днепропетровский национальный университет железнодорожного транспорта имени академика В. Лазаряна, г. Днепропетровск, Украина

Мультиагентная адаптация гибридного генетического алгоритма для обучения нейросетей

Предложен агентно-ориентированный подход адаптации формирования структуры и обучения нейросети к обучающей выборке. Для адаптации структур нейронных сетей используется генетический алгоритм с вещественным кодированием хромосом. Обучение нейросетей выполняется гибридным генетическим алгоритмом с градиентным дообучением лидера. Для подбора параметров обучения используются интеллектуальные агенты, система знаний которых построена по принципу «начальник – подчиненный». Построение знаний осуществляется методом кластеризации. Организация вычислительного процесса позволяет выполнять распределённые вычисления в гетерогенных локальных сетях.

Введение

Аппарат нейронных сетей (НС) позволяет решать широкий круг задач классификации и регрессии [1-3], при этом ключевыми факторами, влияющими на скорость и точность нахождения решения (временную и функциональную эффективность [4], [5]), выступают удачно подобранная структура НС и алгоритм обучения. Таким образом, выбор метода решения задачи диктуется самой задачей. Так, например, при решении регрессионных задач характер воспроизводимой НС функции, которая задаётся обучающей выборкой (ОВ), будет диктовать выбор структуры НС и алгоритма её обучения, потому как нет гарантии, что НС с произвольно выбранными структурой и алгоритмом обучения, смогут воспроизвести функцию с достаточной точностью.

Данная работа является продолжением исследования возможности моделирования рабочей нагрузки СУБД [6]. Именно поэтому в работе рассматриваются прямоточные многослойные НС с сигмовидной активационной функцией, как наиболее приспособленные для решения задач экстраполяции [1]. До сих пор в теории построения нейронных сетей нет окончательно сформулированных правил выбора структуры НС и применения определённого алгоритма обучения. Очевидно, что функциональная эффективность пары <структура НС>+<алгоритм обучения> зависит от параметров её составляющих. Таким образом, для эффективного решения задач нейросетевого моделирования необходимо адаптировать параметры пары <структура НС>+<алгоритм обучения> к особенностям обучающей выборки.

Альтернативой традиционным градиентным алгоритмам обучения НС в настоящее время является использование эволюционных вычислений и, в частности, генетического алгоритма (ГА) [7-10]. В [11] показана эффективность гибридного генетического алгоритма (ГГА) с градиентным обучением лидера.

Применение агентных технологий [12] позволяет распределять частные вычислительные процессы в локальной сети, вменив в обязанности агентов распределение вычислений, сами вычисления и контроль над ними. Тем самым значительно ускоряется общий вычислительный процесс. В данной работе рассматривается мультиагентная эволюционная среда адаптации структуры НС и гибридного генетического обучающего алгоритма к заданной ОВ.

Постановка задачи

Имеется некоторая обучающая выборка. Необходимо подобрать структуру НС и параметры алгоритма обучения такие, чтобы обученная НС имела неизбыточную структуру, воспроизводила ОВ с точностью не менее заданной, причём процесс обучения найденной НС должен занимать минимально возможное время.

Таким образом, необходимо организовать вычислительный процесс, обеспечивающий адаптацию структуры многослойной НС прямого распространения и ГГА к заданной обучающей выборке по критерию.

Основные особенности адаптации обучения нейросетей

Для решения поставленной задачи разработан программный комплекс, решающий три основные задачи:

– адаптация структуры НС;
– адаптация параметров обучающего алгоритма;
– обучение НС заданной структуры.

Для адаптации параметров структуры НС и обучающего алгоритма используется генетический алгоритм параметрической адаптации (ГАПА). Структура НС представляется в виде вектора, элементами которого является количество нейронов в слое. Такое представление обусловлено тем, что в фокусе исследования находятся однотипные прямоточные многослойные нейросети. Строится начальная популяция, в состав которой входят различные структуры: избыточные, минимальные, случайные. Каждой структуре из текущей популяции сопоставляются параметры обучающего алгоритма. Исходя из результатов обучения, для каждой пары <структура НС>+<параметры обучающего алгоритма> рассчитывается эвристический критерий адаптации. Применяется компромиссная оптимизация. Критерий качества адаптации учитывает как качество нейросети (погрешность обучения и размер), так и качество процесса обучения (время обучения и количество итераций алгоритма обучения):

$$E = \begin{cases} 0,2 * \overline{E}_{NN} + 0,3 * \overline{W}_{NN} + 0,4 * \overline{I} + 0,1 * \overline{T}, & если \quad E_{NN} < 2 * \max E_{NN} \\ 1, \quad если \quad E_{NN} \geq 2 * \max E_{NN} \end{cases},$$

где $\max E_{NN}$ – допустимая погрешность НС; E_{NN} – погрешность оцениваемой НС; $\overline{E}_{NN}$ – нормированная погрешность оцениваемой НС; $\overline{W}_{NN}$ – нормированная мощность сети (количество нейронов); $\overline{I}$ – нормированное количество итераций алгоритма обучения НС; $\overline{T}$ – нормированное время обучения НС. Факты о достигнутом критерии адаптации при определённом наборе параметров алгоритма обучения откладываются в хранилище с целью последующего извлечения знаний. Эволюционный процесс продолжается заданное число итераций. Опыт показывает, что во многих случаях достаточно 70 итераций для вырождения популяции до одного вида структуры.

Обучение НС заданной структуры производится гибридным генетическим алгоритмом с градиентным обучением лидера (ГГАО). Перед применением ГГАО для обучения НС необходимо задать взаимнооднозначное соответствие между НС и хромосомой. Для этого используется метод вещественного генетического кодирования (RGA – real genetic coding), при котором отдельными генами хромосом являются вещественные числа. В результате такого кодирования НС преобразуется в хромосому, представляющую собой вектор весовых коэффициентов. Набор из N таких хромосом составляет

начальную популяцию. Данная популяция подвергается стандартным эволюционным операциям (адаптированным к RGA): селекции, скрещиванию и мутации. После очередной итерации эволюционной составляющей ГГАО лучшие НС в популяции дополнительно дообучаются методом обратного распространения ошибки. Процесс продолжается до тех пор, пока лидер популяции не достигнет заданной точности или не будет выполнено максимально возможное число итераций.

Параметрическая адаптация алгоритма обучения

При использовании ГГАО адаптации подлежат параметры как самого генетического алгоритма, так и вспомогательного градиентного метода, применяемого для дообучения лидеров.

Параметры генетического алгоритма можно разбить на четыре группы: параметры мутации, скрещивания, формирования нового поколения, инициализации и обработки популяции.

Для мутации выделено два параметра: вероятность мутации хромосомы в популяции и вероятность мутации определённого гена в хромосоме.

Скрещивание параметризуется вероятностью участия хромосомы в скрещивании, а также количеством точек разрыва (одноточечный, двухточечный, многоточечный кроссовер).

Для формирования нового поколения выделено несколько параметров. Во-первых, – это степень элитизма, которая характеризуется количеством лучших особей, переходящих в новое поколение без изменений и вытесняющих собой худшие особи. Во-вторых, – это функция изменения размера популяции при переходе от старого поколения к новому, то есть в начале эволюционного процесса популяция может намеренно увеличиваться с целью расширения области поиска, а с течением времени – уменьшаться для локализации решения. В-третьих, – это метод отбора хромосом в новое поколение: турнирный, пропорциональный, рулеточный, отбор усечением.

Параметры инициализации популяции – это начальный размер популяции и способ формирования начальной популяции (случайное заполнение хромосом, добавление шума к найденному решению, эвристическое заполнение). Параметры обработки популяции – это максимальное количество итераций ГА и максимальное количество итераций без смены лидера (для реинициализации).

Для градиентной составляющей ГГАО (в работе использовался метод обратного распространения ошибки) выделено три параметра. Во-первых, величина шага, на который будут изменяться весовые коэффициенты сети, во-вторых – количество лидеров, обучаемых градиентным методом, в-третьих – количество проходов градиентного метода при обучении лидера.

Первоначально перечисленные выше параметры формируются случайным образом. Наборы параметров и значений критерия адаптации, достигнутые при этих параметрах, компонуются в кортежи фактов. Адаптация параметров алгоритма обучения происходит на этапе извлечения знаний из накопленных фактов.

Мультиагентная среда адаптации

Работы в сфере так называемых *агентных технологий* рассматривают как отдельно взятые интеллектуальные программные агенты (ИПА), так и мультиагентные системы, и вообще говоря, новый агентно-ориентированный подход к программированию, как раз-

витие объектно-ориентированного программирования. Не существует строгого определения ИПА: различные источники приписывают ИПА разные свойства [5], однако все они сходятся в том, что ИПА – это некий программный объект (в терминах ООП), обладающий *знаниями*, *целью* и *способностью к взаимодействию* как с пользователем-человеком, так и с другими агентами. Мультиагентная система – это среда функционирования и взаимодействия разнородных агентов, деятельность которых подчинена достижению некоторой общей цели, причём эта цель может быть неочевидна для каждого отдельно взятого агента.

В данной работе для решения поставленной задачи была реализована мультиагентная среда адаптации (МСА). В этой среде взаимодействуют девять типов агентов (табл. 1), в результате чего находится единое решение задачи.

Таблица 1 – Агенты МСА

Агент	Ответственность	Цель	Знания
A_1	Агент, формирующий структуры НС (NNarchitector)	Подобрать эффективную структуру НС и параметры её обучения для конкретной задачи	Качество обучения НС определённой структуры при различных сочетаниях параметров обучения
A_2	Агент, обучающий НС заданной структуры (aTeachNN)	Обучить НС определённой структуры алгоритмом обучения с заданными параметрами	Применение гибридного генетического алгоритма для обучения НС
A_3	Агент, распределяющий НС на обучение (dispatchAgent)	Распределить на обучение заданный набор пар <структура НС, параметры обучения>	откуда получать задания на обучение, кому и как их раздавать
A_4	Агент, подбирающий параметры ГА (aControlGA)	Подобрать эффективные параметры обучения ГА для заданной обучающей выборки	Качество обучения при различных параметрах ГА
A_5	Агент, подбирающий параметры градиентного метода (aControlGradient)	Подобрать эффективные параметры обучения градиентным методом для заданной обучающей выборки	Качество обучения при различных параметрах градиентного метода
A_6	Агент, ответственный за параметры мутации (aMutation)	Подобрать эффективные параметры мутации при обучении ГА	Качество обучения при различных параметрах мутации
A_7	Агент, ответственный за параметры скрещивания (aCrossover)	Подобрать эффективные параметры скрещивания при обучении ГА	Качество обучения при различных параметрах кроссовера
A_8	Агент, ответственный за параметры обработки популяции (aPopulation)	Подобрать эффективные параметры формирования популяции при обучении ГА	Качество обучения при различных параметрах формирования популяции
A_9	Агент, ответственный за параметры формирования нового поколения (aNewGeneration)	Подобрать эффективные параметры формирования нового поколения при обучении ГА	Качество обучения при различных параметрах формирования нового поколения

Формирование базы знаний агентов

Интеллектуальная составляющая системы формирования структуры и обучения нейросети заключается в извлечении, накоплении и использовании знаний агентами. Система знаний агентов построена по принципу «начальник – подчиненный». Начальник знает общие задачи и цели и владеет общими знаниями, подчиненные – решают узкие, конкретные задачи и обладают ограниченными знаниями. Принятие решений является

некоторым компромиссом между начальником и подчиненными. В предлагаемой модели один из агентов (A_4) обладает общими знаниями, другие ($A_6 \dots A_9$) – частными. Агент A_5 работает по традиционной технологии, обладает всеми знаниями, необходимыми для выполнения своей ответственности.

Покажем, каким образом формируются знания агентов, затем – как используются.

Перед агентами $A_6 \dots A_9$ стоит задача определить значения некоторого набора параметров $\{P_i\}$, при которых функция цели S будет близка к оптимальной, и сообщить их агенту A_4. При этом каждый агент отвечает за свой набор параметров. Агент A_4, используя поступившие предложения и собственные знания, должен принять решение, какие параметры в дальнейшем использовать. Так как значения параметров «подчиненных» могут отличаться от соответствующих значений «начальника» для обоснованного принятия решений значения параметров снабжаются атрибутом «степень уверенности в используемых знаниях» или «уровень доверия», обозначим его ν. Показатель ν соответствует каждому набору параметров $\{P_i\}$.

Первоначально агенты начинают работать в условиях отсутствия знаний, и некоторое время занимаются только сбором информации. При этом $\nu = 0$.

Агент экспериментирует – случайным образом определяет значения параметров P_{ij} (i-го параметра в j-м опыте). В результате эксперимента агент получает значения параметров P_{ij} и значения критерия качества полученного результата, в нашем случае компромиссного критерия K_j. Агент $A_6 \dots A_9$ накапливает таблицу фактов в виде множества кортежей параметров и соответствующего критерия качества $F = \{f_j\} = \{(P_1, P_2, \dots, P_m, K_j)\}$.

При наличии достаточного объема накопленной информации агент периодически начинает извлекать и постоянно использовать знания (через каждые 50 фактов).

Первоначально все параметры нормализуются. Полученное множество $\overline{F}$ разбивается на несколько частей $\overline{F}_i$ ($N = 10$), соответствующих определенным уровням качества:

$$f_j \subset \overline{F}_i \mid (i-1) \cdot (\max(K_j) - \min(K_j)) / N < K_j < i \cdot (\max(K_j) - \min(K_j)).$$

Каждое множество $\overline{F}_i$ кластеризуется методом максиминного расстояния [14]. Ищется множество центров кластеров I. Находятся два максимально удаленных друг от друга кортежа в m-мерном Евклидовом пространстве. Они принимаются за центры двух начальных кластеров. Далее последовательно ищутся точки:

$$\bar{f}_k \subset \overline{F}_s \mid \min_{\bar{f}_i \in I} (\rho(\bar{f}_i, \bar{f}_k) = \max_{\bar{f}_j \in \overline{F}_s} (\min_{\bar{f}_i \in I} (\rho(\bar{f}_i, \bar{f}_j)),$$

где $\bar{f}_i$ – точка с координатами $(\overline{P}_1, \overline{P}_2, \dots, \overline{P}_m)$ – нормированными значениями параметров. Если $\min\limits_{\bar{f}_i \in I} (\rho(\bar{f}_i, \bar{f}_k) \geq \dfrac{1}{6} \sum\limits_{\bar{f}_i, \bar{f}_j \in I} \rho(\bar{f}_i, \bar{f}_j)$, то точка $\bar{f}_k$ принимается за новый центр кластера. Таким образом, формируется множество центров кластеров, остальные точки разносятся к кластерам по критерию минимума расстояния.

Для дальнейших расчетов выбирается точка с $\min K_j$ из лучшего кластера (по среднему K_j для точек кластера) из множества $\overline{F}_1$.

Если области кластеров первого уровня и кластеров уровней $2 \dots N$ не пересекаются, то можно утверждать, что области кластеров первого уровня и являются областями, соответствующими наилучшим параметрам. При этом уровень доверия к полученным данным максимальный, ввиду отсутствия противоречий. Если же указанные области пересекаются, то уровень доверия, конечно же, ниже.

Представим ситуацию, когда области наилучшего и наихудшего кластеров совпадают, наилучшие и наихудшие параметры перемежаются. Естественно уровень доверия к таким данным как к наилучшим очень небольшой.

Из приведенных соображений уровень доверия к параметрам вычисляется как:

$$\nu = \sum_{f_j \in \bar{F}} a_j \Big/ \sum_{f_j \in \bar{F}} b_j ,$$

где $a_j = \begin{cases} 1 - 0{,}1(i-1), \text{если } \rho(\bar{f}_k, \bar{f}_j) = \min\limits_{\bar{f}_s \in I,\, \bar{f}_s \in \bar{F}_i} (\rho(\bar{f}_s, \bar{f}_j)) \\ 0, \text{в противном случае} \end{cases}$, $b_j = \begin{cases} 1, \text{если } a_j > 0 \\ 0, \text{в противном случае} \end{cases}$.

Численный эксперимент

С помощью МСА был проведён следующий вычислительный эксперимент. В качестве обучающих выборок были выбраны такие функции:

– $f(x, y) = x \text{ xor } y$, где $x \in [0,1]$, $y \in [0,1]$;
– $f(x, y) = \sin(x) * y$, где $x = 0.1, 0.2 \ldots 3$, $y = 3, 2.9, \ldots 0.1$;
– $f(x, y) = x * x * y$, где $x = 0.1, 0.2 \ldots 3$, $y = 3, 2.9, \ldots 0.1$.

Обучение каждой функции происходило при четырёх различных начальных условиях: без знаний о параметрах алгоритма обучения и со знаниями о параметрах алгоритма обучения, построенных на фактах, полученных при обучении трём вышеуказанным функциям. Допустимая погрешность при обучении НС определялась как среднеквадратичное отклонение и задавалось равным $\max E_{NN} = 0{,}005$. Максимальное количество итераций ГАПА устанавливалось равным 70. Результаты проведённого эксперимента представлены в табл. 2.

Таблица 2 – Результаты адаптации обучения нейросети

Функция	Показатель	Без знаний	Со знаниями		
			XOR	SIN(X)*Y	X*X*Y
XOR	Погрешность	0,0043	0,0017	0,0047	0,0045
	Размер сети	2:8:1	2:2:1	2:2:1	2:7:1
	Обучение (итер. ГГАО / мсек)	8 / 31	166 / 4469	58 / 203	57 / 313
	Итерации ГАПА	69	4	7	14
SIN(X)*Y	Погрешность	0,0031	0,0048	0,0037	0,0049
	Размер сети	2:3:1	2:4:1	2:2:1	2:4:1
	Обучение (итер. ГГАО / мсек)	4 / 16	142/ 1641	604 / 10109	545 / 6078
	Итерации ГАПА	69	40	9	16
X*X*Y	Погрешность	0,0074	0,0043	0,0056	0,0049
	Размер сети	2:8:1	2:1:1	2:3:1	2:3:1
	Обучение (итер. ГГАО / мсек)	3013 / 125563	1124 / 41437	527 / 7435	331 / 1953
	Итерации ГАПА	39	33	4	15

Так как в критерий качества адаптации заложена мощность нейросети, задаваемая количеством нейронов, то в процессе итераций ГАПА среднее в популяции количество нейронов в сети уменьшается. На рис. 1 показана регрессия среднего количества слоёв нейросетей на количестве итераций ГАПА, из которого видно, что с использованием знаний структура НС упрощается быстрее.

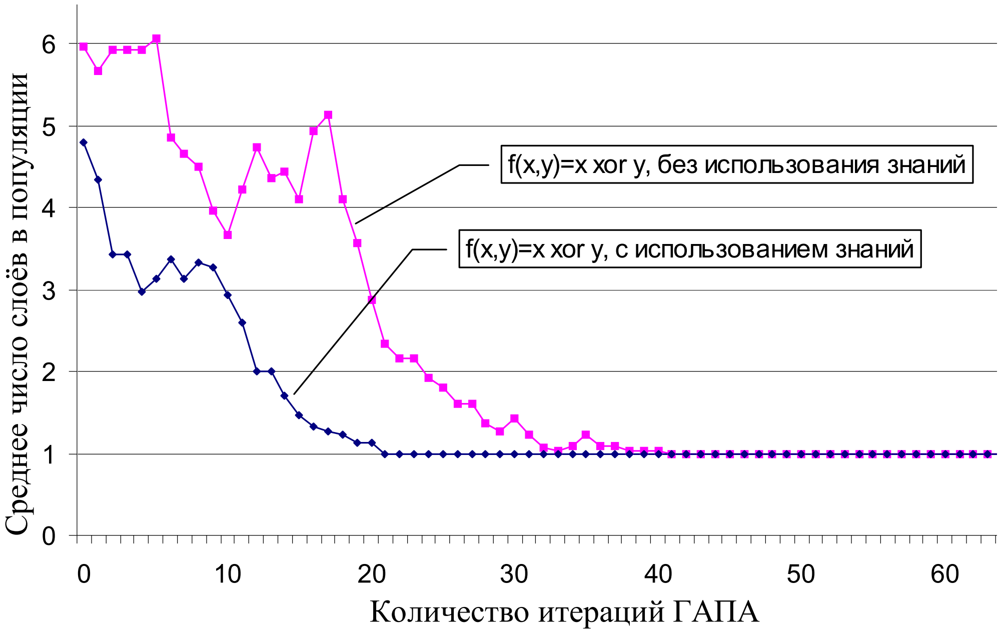


Рисунок 1 – Зависимость среднего количества слоёв от итераций ГАПА

Заключение

Разработан и апробирован программный комплекс формирования структуры и обучения прямоточных многослойных нейросетей. Разработанная методика обучения нейросетей обладает высоким уровнем интеллектуализации и позволяет не только формировать нейросеть, но и переоидически её перестраивать. Одним из недостатков нейросетей является то, что их поведение адекватно моделируемым объектами процесссам в области обучения, которую к тому же сложно как-то определить и описать. Предлагаемые возможности дообучения позволяют повышать адекватность нейросети в условиях изменяющихся моделируемых процессов.

Особенность программного комплекса заключается в организации вычислительного процесса на основе агентной технологии. Агенты могут рассылаться по локальной сети и выполнять поставленные им задачи по обучению нейросети. Вычисления могут выполняться в гетерогенной локальной сети и с максимальным использованием вычислительных ресурсов многоядерных процессоров.

Показана целесообразность адаптации алгоритма обучения. При переобучении нейросети (если она была обучена близким функциям) процесс обучения ускоряется и в результате достигаются лучшие структуры нейросети.

Литература

1. Нейронные сети. STATISTICA Neural Networks : Пер. с англ. – М.: Горячая линия –Телеком, 2000. – 182 с.
2. Заенцев И.В. Нейронные сети: основные модели. – Воронеж: Воронежский государственный университет, 1999. – 76 с.
3. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. – М.: Горячая линия – Телеком, 2006 . – 452 с.
4. Шинкаренко В.И. Функциональная эффективность нечетко специфицированных алгоритмов // Проблемы программирования. – 2006. – № 1. – С. 24-33.
5. Шинкаренко В.И. Сравнительный анализ временной эффективности функционально эквивалентных алгоритмов // Проблемы программирования. – 2001. – № 3-4. – С. 31-39.

6. Шинкаренко В.И., Олейник Д.В. Нейросетевое моделирование рабочей нагрузки СУБД // Искусственный интеллект. – 2007. – № 4. – С. 657-664.
7. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. – М.: Физматлит, 2006. – 320 с.
8. Емельянов В.В., Курейчик В.М., Курейчик В.В. Теория и практика эволюционного моделирования. – М.:ФИЗМАТЛИТ, 2005. – 432 с.
9. Цой Ю.Р., Спицын В.Г. Использование генетического алгоритма для настройки весов и структуры искусственной нейронной сети // Молодежь и современные информационные технологии. Сборник трудов II Всероссийской научно-практической конференции студентов. – Томск: Изд-во ТПУ. – 2004. – С. 221-223.
10. Цой Ю.Р., Спицын В.Г. Генетический алгоритм настройки искусственной нейронной сети // Тезисы докладов конференции конкурса студентов, аспирантов и молодых ученых «Технологии Microsoft в информатике и программировании». – Новосибирск: НГУ. – 2004. – С. 131-133.
11. Паклин Н.Б., Сенилов М.А., Тененев В.А. Интеллектуальные модели на основе гибридного генетического алгоритма с градиентным обучением лидера // Искусственный интеллект. – 2004. – № 4. – С. 159-168.
12. Плескач В.Л., Рогушина Ю.В. Агентні технології: Монографія. – К.: Київ. нац. торг.-екон. ун-т, 2005. – 338 с.
13. Атанасова Т. Агентная технология: концепции, модели, приложения. – Варна, 2000. – 155 с.
14. Ту Дж., Гонсалес Р. Принципы распознавания образов. – 1978. – 411 с.

Олійник Д.В., Шинкаренко В.І.

Мультіагентна адаптація гібридного генетичного алгоритму для навчання нейромереж

Запропонований агентно-орієнтований підхід адаптації формування структури та навчання нейромережі до навчальної вибірки. Для адаптації структур нейронних мереж використовується генетичний алгоритм з кодуванням хромосом в дійсних числах. Навчання нейромереж відбувається гібридним генетичним алгоритмом з градієнтним донавчанням лідера. Для підбору параметрів навчання використовуються інтелектуальні агенти, система знань яких побудована по принципу «начальник – підлеглий». Побудова знань відбувається методом кластеризації. Организація обчислювального процесу дозволяє виконувати розподіленні обчислення в гетерогенних локальних мережах.

D.V. Oleynik, V.I. Shinkarenko

Multiagent Adaptation of Hybrid Genetic Algorithm for Neuronets Learning

The agent-oriented method for adaptation of forming and learning of neuronet to learning selection is suggested. Genetic algorithm with real genetic coding is used for adaptation of neuronet structure. Neuronets learning is performing by means of hybrid genetic algorithm with gradient leader relearning. The intellectual agents are used for obtaining of learning parameters. Their knowledge system is based on “chief-inferior”. Knowledge building is performed by means of clusterization. Organization of calculating process allows to perform distributed calculations in heterogeneous local area networks.

Статья поступила в редакцию 18.07.2008